The ABCs of centromeres

Barbara Mellone, Sylvia Erhardt and Gary H. Karpen

Accurate segregation of genetic information during cell division relies on a multiprotein complex called the kinetochore, whose formation requires specialized centromeric chromatin. Two papers in this issue of *Nature Cell Biology* identify a multitude of new vertebrate kinetochore proteins that provide insight into the link between centromeric chromatin and the kinetochore, and suggest a functional relationship between centromeres and nucleoli during interphase.

Kinetochores — a specialized structure that forms at centromeres — serve as the contact point between spindle microtubules and the chromosomes, ensuring normal segregation to daughter cells during mitosis and meiosis. In principle, this function could be accomplished by a single protein that binds both centromeric chromatin and microtubules, coupled with regulation of microtubule assembly and disassembly. However, kinetochores are required for other processes, such as mediating the spindle assembly or mitotic checkpoint[1], suggesting the presence of multiple proteins and subcomplexes devoted to accomplishing these different roles.

The first vertebrate centromere proteins (CENPs A–C) were identified by their reactivity with sera from human patients with unusual autoimmune disorders. Further studies ultimately extended this 'alphabet soup' of proteins to CENP-I, and elucidated the spatial organization and functions of centromere and kinetochore proteins. Many of these components and structures are conserved across eukaryotes, suggesting that common mechanisms regulate kinetochore formation and activities.

The sites of kinetochore formation are determined by centromeric DNA and chromatin, whose identity and propagation seem to be regulated by epigenetic mechanisms in most eukaryotes (for a review, see ref. 2). Centromeric chromatin is distinguished by the presence of the histone H3 variant CENP-A, which is essential for kinetochore formation and function, and may serve as an epigenetic mark that propagates centromere identity through replication and cell division[2]. In humans and flies, interphase centromeric chromatin is characterized by interspersed regions of CENP-A-containing and histone H3-containing nucleosomes, and the H3 nucleosomes contain post-translational modifications that are distinct from those observed in canonical euchromatin and heterochromatin[3,4]. In mitotic chromosomes, CENP-A and H3 nucleosome blocks are organized into spatially distinct domains to create a polarized structure that serves as a foundation for kinetochore formation[3]. Moving poleward from centromeric chromatin, previous studies identified centromere-associated proteins such as CENP-C, -H and -I, as well as outer kinetochore proteins that are present only during mitosis but are required for spindle attachment, chromosome movements and the mitotic checkpoint.

Despite the identification of numerous centromere and kinetochore proteins, many questions about the assembly, organization and functions of these structures remain unanswered. For example, we still lack a clear understanding of how CENP-A is exclusively deposited at centromeres, and how centromeric chromatin is linked to the outer kinetochore[5]. On pages 458 and 446 of this issue, Foltz *et al.*[6] and Okada *et al.*[7] describe the identification and analysis of a surprisingly large group of vertebrate kinetochore proteins (CENP K–U) that add to our knowledge of these processes. To identify proteins associated specifically with centromeric chromatin, Foltz *et al.* compared complexes associated with nucleosomes containing the histone variants H3.1 and CENP-A by tandem affinity purification (TAP) and mass spectrometry analysis[6]. In a parallel study, Okada *et al.* identified complexes associated with the chromatin-associated proteins CENP-H and CENP-I in human and chicken cells[7].

The two independent studies identified overlapping sets of factors, including novel and known kinetochore components[8]. Four known centromeric proteins (CENP-B, -C, -H and –U) and three novel proteins (CENP-M, -N and -T) were isolated by their association with one or a few CENP-A nucleosomes. Foltz *et al.* have named this complex CENP-A NAC (nucleosome associated complex). TAP tagging of NAC components (CENP-M, -N and -U) led to the identification of six additional components: CENP-I, -K, -L, -O, -P, -Q, -R and -S. These proteins were not identified in the CENP-A–TAP purifications, suggesting that they form a complex distinct from centromeric chromatin and NAC proteins (Fig. 1). Because of this biochemical distinction, this complex was named CENP-A CAD (for CENP-A distal). Verification of the structural organization suggested by the biochemical analysis of these complexes requires further analysis.

Homologues for CENP-M through -R were identified in some vertebrate organisms, but orthologues have not been detected by homology in more distant eukaryotes, such as flies and worms. Protein domain analysis has not yet suggested potential biochemical functions for the NAC and CAD complexes. However, components of the NAC and CAD complexes are constitutively associated with the vertebrate centromere, suggesting that they may establish a bridge between centromeric chromatin and

Barbara Mellone, Sylvia Erhardt and Gary H. Karpen are in the Department of Genome Biology, Lawrence Berkeley National Lab, and Department of Molecular and Cell Biology, UC Berkeley, Berkeley, CA 94720, USA.
e-mail: karpen@fruitfly.org

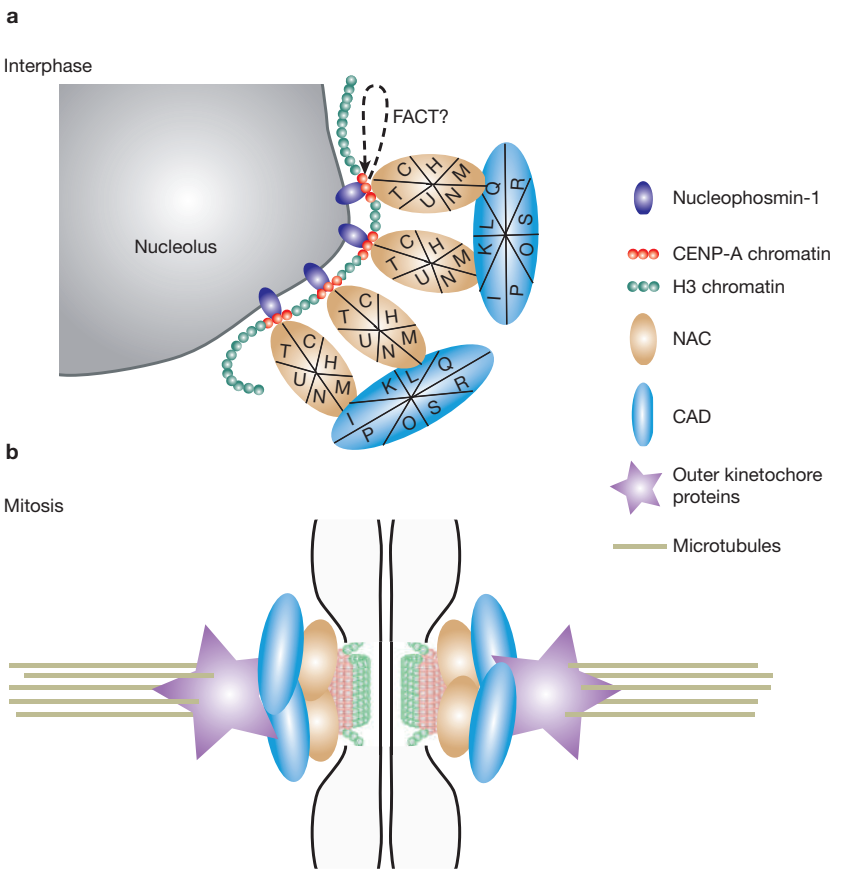


Figure 1 A schematic representation of two hypothetical states of the centromere. One state is 'protected' by its association with the nucleolus during interphase and the other is 'competent' for kinetochore formation in mitosis. (**a**) CENP-A is responsible for recruiting the NAC complex, which in turn recruits the CAD complex. Centromeric chromatin may be anchored to the nucleolus through an interaction with nucleophosmin-1; this association could be important for centromere structure or function. FACT remodelling may promote CENP-A chromatin assembly or function. (**b**) At the onset of mitosis, the centromere could be released by nucleolar disassembly, which allows association with other mitotic kinetochore-associated factors and kinetochore formation.

kinetochore formation during mitosis, and/or regulate CENP-A deposition or maintenance.

These possibilities were tested by both groups through analysis of the effects of protein depletion on the function and localization of kinetochore components. In general, NAC components are required to recruit CAD complex proteins, which in turn recruit subsets of outer kinetochore proteins. Depletion of individual CENPs caused increased mitotic indices, lagging and misaligned chromosomes, and, in some cases, hypercondensed chromosomes. In addition, NAC and CAD complex components are important for achieving and sustaining stable kinetochore–microtubule interactions, but not for checkpoint signalling. Notably, hMis12 and the outer kinetochore components Hec1–ndc80, CENP-E and CENP-F are properly localized in cells lacking the NAC component CENP-U(50). This indicates that outer kinetochore assembly involves independent 'modules', and that defects in microtubule–kinetochore interactions and chromosome segregation are not due to complete disruption of the outer kinetochore. Nevertheless, these results suggest that, in general, NAC components are required for recruitment of CAD and some outer kinetochore proteins, favouring the idea that these new constitutive centromere components provide a 'missing link' for the assembly of kinetochore subcomplexes during mitosis.

These studies also provide insights into factors involved in CENP-A assembly onto centromeric chromatin. Previous work suggested a role for a chromatin assembly factor (CAF-1) subunit (p55–Mis16–RbAp46–48) in CENP-A loading at the centromere[9]. Foltz *et al.* observed that chromatin containing H3.1 nucleosomes is associated with this CAF-1 component[6], and other studies have shown that replication-independent assembly of the H3 variant H3.3 involves a different complex called HIRA[10]. However, Foltz *et al.* did not observe associations of CAF-1 or HIRA components with CENP-A chromatin: the absence of HIRA components is somewhat surprising, given that CENP-A nucleosomes are also assembled independently of DNA replication[11]. It is possible that CAF-1 and/or HIRA are required for initial assembly of CENP-A chromatin and subsequently disassociate from CENP-A chromatin. However, this would mean that the same complexes display differential retention after assembly of H3 versus CENP-A nucleosomes. The authors suggest that it is more likely that CENP-A assembly is mediated by complexes distinct from both CAF-1 and HIRA[11]. Finally, FACTp140 and FACTp80, which are involved in chromatin remodelling and transcription, were isolated specifically in CENP-A–TAP purifications, similarly to previous findings[8]. This intriguing interaction suggests that transcription may pass through CENP-A chromatin, or that chromatin remodelling has a role in CENP-A nucleosomes formation and/or maintenance. It will be very interesting to determine whether FACT is required for centromere formation or function.

Okada *et al.* provide data supporting the idea that NAC components are required for assembly of CENP-A chromatin[7]. They demonstrate that knockouts of chicken CENP-H, -I, -K, -L and -M block centromeric incorporation of CENP-A–GFP expressed from a constitutive promoter, although localization of endogenous CENP-A was unaffected. Previous studies showed that the *Schizosaccharomyces pombe* CENP-I homologue (Mis6) is required for centromere localization of CENP-A (Cnp1)[12], but chicken CENP-I knockouts did not display mislocalization of endogenous CENP-A[13]. The results of Okada *et al.* now suggest that CENP-I and associated proteins are required for incorporation of newly-synthesized, but not previously assembled, CENP-A nucleosomes, thus providing an explanation for these previously contradictory findings. Although the suggestion that NAC components are required for proper CENP-A localization is exciting, NAC recruitment to centromeres is reciprocally dependent on CENP-A. Thus, the key molecular mechanisms and components responsible for

specifying and propagating CENP-A localization and centromere identity remain elusive.

Foltz *et al.* make the intriguing observation that the nucleolar protein, nucleophosmin-1, is associated with CENP-A chromatin. This is interesting in light of previous observations that centromeres are clustered around nucleoli during interphase in fly and human cells, that centromere proteins are present in purified nucleoli[14] and that a nucleolar transcription factor interacts with CENP-C[15]. Previous studies demonstrated that enhancer-blocking insulators are also associated with nucleophosmin[16]. It is possible that nucleolus 'anchoring' is a conserved mechanism for sequestering centromeres and other specialized chromatin sites (Fig. 1). This interaction may be important for centromeric chromatin assembly, higher order structure, or promoting or reducing accessibility to other factors. Additional work is needed to clarify the relationship between nucleoli and centromeres. Determining how important these observations are requires direct assessment of the effects of depleting nucleolar proteins on the composition and function of centromeric chromatin, and subsequent kinetochore formation during mitosis.

These studies have identified a large number of new proteins associated with vertebrate centromeric chromatin, and demonstrated their importance to CENP-A incorporation, kinetochore formation and chromosome segregation. Future studies based on these results, and other intriguing observations, are likely to generate a more complete understanding of the spatial organization and functions of these complexes, as well as the molecular mechanisms involved in centromere identity, propagation and kinetochore assembly. However, these studies do bring us perilously close to the end of the alphabet, should future studies identify additional CENPs. □

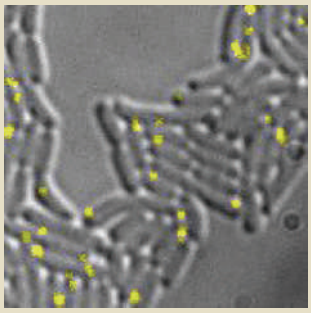
1. Cleveland, D.W., Mao, Y. & Sullivan, K.F. *Cell* **112,** 407–421 (2003).
2. Sullivan, K.F. *Curr. Opin. Genet. Dev.* **11,** 182–188 (2001).
3. Blower, M.D., Sullivan, B.A. & Karpen, G.H. *Dev. Cell* **2,** 319–330 (2002).
4. Sullivan, B.A. & Karpen, G.H. *Nature Struct. Mol. Biol.* **11,** 1076–1083 (2004).
5. Mellone, B.G. & Allshire, R.C. *Curr. Opin. Genet. Dev.* **13,** 191–198 (2003).
6. Foltz, D.R. *et al. Nature Cell Biol.* **8,** 458–469 (2006).
7. Okada, M. *et al. Nature Cell Biol.* **8,** 446–457 (2006).
8. Obuse, C. *et al. Genes Cells* **9,** 105–120 (2004).
9. Chen, E.S., Saitoh, S., Yanagida, M. & Takahashi, K. *Mol. Cell* **11,** 175–187 (2003).
10. Ray-Gallet, D. *et al. Mol. Cell* **9,** 1091–1100 (2002).
11. Shelby, R.D., Monier, K. & Sullivan, K.F. *J. Cell Biol.* **151,** 1113–1118 (2000).
12. Takahashi, K., Chen, E.S. & Yanagida, M. *Science* **288,** 2215–2219 (2000).
13. Nishihashi, A. *et al. Dev. Cell* **2,** 463–476 (2002).
14. Ochs, R.L. & Press, R.I. *Exp. Cell Res.* **200,** 339–350 (1992).
15. Pluta, A.F. & Earnshaw, W.C. *J. Biol. Chem.* **271,** 18767–18774 (1996).
16. Yusufzai, T.M., Tagami, H., Nakatani, Y. & Felsenfeld, G. *Mol. Cell* **13,** 291–298 (2004).

Protein expression: one by one

Most of what we assume to be true about gene expression is based on genetic and biochemical studies on total pools of molecules and cells, and even single-cell measurements have so far lacked the sensitivity to allow observation of protein expression at the single-molecule level. Now Sunney Xie and colleagues describe two powerful techniques that can track single protein expression, even of low-copy number proteins.

The technique reported in *Nature* (**440,** 358–362; 2006) is based on the 'veteran' gene reporter β-galactosidase (β-gal) that is expressed from the *lacZ* gene. Although β-gal is a highly-sensitive probe, the fluorescent molecules it produces, following substrate hydrolysis, are not retained in the cell. The authors used closed microfluidic chambers to trap the fluorescent molecules excreted by the cells in the small volume of the chambers. In doing so, they were able to obtain real-time quantitative information on gene expression in live *Escherichia coli* cells with single molecule sensitivity. Furthermore, they showed that this technique was also applicable to budding yeast and mouse embryonic stem cells expressing β-gal from the *GAL1* or *ROSA* promoters, respectively.

The second technique, reported in *Science* (**311,** 1600–1603; 2006), replaces the native *lacZ* gene with a fusion protein of a fluorescent tag (YFP–Venus) and Tsr (a membrane protein), so it can be used as a reporter for monitoring protein expression from the *lac* promoter. By tracking the disappearance of the fluorescence signal after photobleaching, the authors could show that each fluorescent peak corresponded to a single molecule.



An overlay of the DIC and fluorescence images of *E. coli* cells expressing the fluorescent protein Venus, tethered to the membrane protein Tsr. Single Tsr–Venus fusion molecules (yellow spots) can be detected when they anchor to the inner membrane of the cell.

In both studies, the authors concluded that protein molecules are produced in bursts randomly occurring over time, that the number of molecules per burst follows an exponential distribution, and that each burst results from a stochastically transcribed single mRNA. Furthermore, the burst size and frequency could be determined either by real-time quantitative monitoring of protein production or by measuring the steady-state distribution of the number of protein copies within a population of cells.

Xie and colleagues have developed two highly related methods that allow single-molecule sensitivity at a single-cell level. These techniques offer new possibilities for understanding gene expression and will allow genome-wide characterization of low-copy number proteins.

MYRTO RAFTOPOULOU